

SEQUENCE LISTING

<110> Gellissen, Gerd
 Braus, Gerhard
 Pries, Ralph
 Krappmann, Sven
 Strasser, Alexander

<120> Nucleic Acid Molecule Comprising a Nucleic Acid Coding for a Polypeptide with Chorismate Mutase Activity

<130> 029474-5007-00

QA
 <140> 10/042059
 <141> 2001-10-25

<150> DE 199 19 124.7
 <151> 1999-04-27

<160> 7

<170> PatentIn version 3.1

<210> 1
 <211> 843
 <212> DNA
 <213> Hansenula polymorpha

<220>
 <221> CDS
 <222> (1)..(843)
 <223>

<400> 1
 atg gac ttt atg aag cca gaa aca gtg ctg gac ctt ggc aac att aga
 Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg
 1 5 10 15

gat gcc ttg gtc cgg atg gag gat acg atc atc ttc aac ttt atc gag
 Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu
 20 25 30

cgg tcg cag ttc tat gcg tcg ccc tcg gta tac aaa gtc aac cag ttc
 Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe
 35 40 45

cct att ccc aac ttc gac ggc tcg ttc ttg gac tgg ctg ttg tcg cag
 Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln
 50 55 60

cac gag cga atc cat tcg cag gtg agg aga tac gac ggc cca gac gag
 His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu
 65 70 75 80

gtg cct ttt ttc ccc aac gtg ctg gaa aaa acg ttt ctg ccc aag atc
 Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile
 85 90 95

RECEIVED
 TECH CENTER 1600/2900
 02 SEP 11 PM 1:56

<210> 2
<211> 280
<212> PRT
<213> Hansenula polymorpha

<400> 2

Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg
1 5 10 15

Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu
20 25 30

Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe
35 40 45

Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln
50 55 60

His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu
65 70 75 80

Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile
85 90 95

Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn
100 105 110

Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala
115 120 125

Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala
130 135 140

Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg
145 150 155 160

Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp
165 170 175

Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn
180 185 190

Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg
195 200 205

Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser
210 215 220

Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile

225

230

235

240

Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu
245 250 255

Asp Glu Glu Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp
260 265 270

Arg Phe Leu Ser Ser Gly Leu Tyr
275 280

<210> 3
<211> 1655
<212> DNA
<213> *Hansenula polymorpha*

<220>
<221> gene
<222> (1)..(1655)
<223> 1,8 kb genomic DNA-fragment from *Hanseula polymorpha*

<220>
<221> gene
<222> (1)..(1655)
<223> 1,8 kb genomic DNA-fragment from *Hansenula polymorpha*

<400> 3
cccgccccaa tgccagcaat atggagacgt ttaggcagaa taggcgttcc atacttctca 60
cgctgcttgt tgccaccgga atatacaccg cattgcagtt tgcacacatc atactatatg 120
acgattacat tggcggaacg tatcgcgagt cgctcacgag acgcattaga atgacagaga 180
aatcgcgaaa cgaccttata gacgcacgtg aaaactacgg gtttggaggc agcaaggagg 240
agcgaatcca gcggttttg tggttcagac atcttcgtg gcttttaggc gaggataagc 300
gaacttgagg agcggtttttt ttttcctgtt tagttttgt aggtatggac tttatgaagc 360
cagaaaacagt gctggacctt ggcaacattha gagatgcctt ggtccggatg gaggataacg 420
tcatcttcaa ctttatcgag cggtcgcagt tctatgcgtc gccctcgta tacaaagtca 480
accagttccc tattccaaac ttgcacggct cgttcttggc ctggctgttgc tcgcagcacg 540
agcgaatcca ttgcaggtg aggagatacg acgcgccaga cgaggtgcct ttttccccca 600
acgtgctgga aaaaacgttt ctgcccaga tcaactaccc atcggtgcta gcctcctacg 660
cgatgaaat caacgtcaac aaagagatac tcaagatcta cacgtcagag atagtaccag 720

| | | | | | | |
|-------------|-------------|------------|--------------|------------|-------------|------|
| gaatagctgc | aggcagcgg | gagcaggagg | acaaccttgg | ctcggtcgca | atggccgaca | 780 |
| tcgagtgcct | gcagtcgcta | tccagaagaa | tccatTTTgg | ccgtttgtc | gcagaggcta | 840 |
| aatttatcag | tgagggggac | aagattgtgg | atctgatcaa | aaagagagat | gtgaaaggca | 900 |
| ttgaggcgct | catcacaaac | gccgaggtcg | aaaaacggat | cttggacaga | cttctggaga | 960 |
| agggaaagggc | gtatggaaca | gacccgacac | taaagttcac | gcagcacatt | cagagcaagg | 1020 |
| tgaagccgga | ggtgattgtg | aaaatctaca | aggatttcgt | gattccgctc | acgaagaagg | 1080 |
| tcgaagtcga | ctacttgctg | agacggctgg | aggacgagga | ggacgatgat | gcgacgcaga | 1140 |
| aaagcggcgg | ctacgttgac | cggTTTCTCT | cctctggctt | gtactagaaa | ttaaaatttt | 1200 |
| cagtacttta | attattctcg | aattctagtt | cagataccgc | atggtaattt | caaaggccag | 1260 |
| aaaagtggcc | gcgttggctg | gggcagctct | cagaatagtc | ggcgagaatc | ctttgactag | 1320 |
| cccccaggca | ccgctctgtc | tccaaatacc | cctaatacgatc | tcaacagcat | ttctataaaac | 1380 |
| cagcttcttg | tagttgtccg | tctgcatgtt | ggacttgate | acatcgatcg | gataaatact | 1440 |
| gaaccacatc | ccgtaacctg | ccagcgcccc | aaagacgcag | agcttccagt | tctcgatgtc | 1500 |
| cttcctggca | atattccgctg | actcgatctc | gttttcacg | agagcttcaa | aagtcagaaa | 1560 |
| atacgctccg | ctacccaaac | tttcttttgc | cagcgttaggt | cccagacccc | ggtagattaa | 1620 |
| cttgatgcct | cccgatgtt | acagcttctt | gatcc | | | 1655 |

<210> 4
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 4
aattaaccct cactaaaggg 20

<210> 5
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 5
gtaatacgac tcactatagg gc 22

<210> 6
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 6
atatacatct acaaaaacta aacagg

26

<210> 7
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 7
atatacatct gatgcgacgc agaaaaagc

28

SEQUENCE LISTING

<110> Gellissen, Gerd
Braus, Gerhard
Pries, Ralph
Krappmann, Sven
Strasser, Alexander

<120> Nucleic Acid Molecule Comprising a Nucleic Acid Coding for a Polypeptide with Chorismate Mutase Activity

<130> 029474-5007-00

<140> 10/042059
<141> 2001-10-25

<150> DE 199 19 124.7
<151> 1999-04-27

<160> 7

<170> PatentIn version 3.1

<210> 1
<211> 843
<212> DNA
<213> Hansenula polymorpha

<220>
<221> CDS
<222> (1)..(843)
<223>

| | |
|---|-----|
| <400> 1 | 48 |
| atg gac ttt atg aag cca gaa aca gtg ctg gac ctt ggc aac att aga | |
| Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg | |
| 1 5 10 15 | |
| gat gcc ttg gtc cgg atg gag gat acg atc atc ttc aac ttt atc gag | 96 |
| Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu | |
| 20 25 30 | |
| cgg tcg cag ttc tat gcg tcg ccc tcg gta tac aaa gtc aac cag ttc | 144 |
| Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe | |
| 35 40 45 | |
| cct att ccc aac ttc gac ggc tcg ttc ttg gac tgg ctg ttg tcg cag | 192 |
| Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln | |
| 50 55 60 | |
| cac gag cga atc cat tcg cag gtg agg aga tac gac gcg cca gac gag | 240 |
| His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu | |
| 65 70 75 80 | |
| gtg cct ttt ttc ccc aac gtg ctg gaa aaa acg ttt ctg ccc aag atc | 288 |
| Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile | |
| 85 90 95 | |
| aac tac cca tcg gtg cta gcc tcc tac gcg gat gaa atc aac gtc aac | 336 |
| Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn | |

| 100 | 105 | 110 | |
|--|-----|-----|-----|
| aaa gag ata ctc aag atc tac acg tca gag ata gta cca gga ata gct Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala | 115 | 120 | 384 |
| | 120 | 125 | |
| gca ggc agc gga gag cag gag gac aac ctt ggc tcg tgc gca atg gcc Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala | 130 | 135 | 432 |
| | 135 | 140 | |
| gac atc gag tgc cag tcg cta tcc aga aga atc cat ttt ggc cgt Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg | 145 | 150 | 480 |
| | 150 | 155 | 160 |
| ttt gtc gca gag gct aaa ttt atc agt gag ggg gac aag att gtg gat Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp | 165 | 170 | 528 |
| | 170 | 175 | |
| ctg atc aaa aag aga gat gtg gaa ggc att gag gcg ctc atc aca aac Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn | 180 | 185 | 576 |
| | 185 | 190 | |
| gcc gag gtc gaa aaa cgg atc ttg gac aga ctt ctg gag aag gga agg Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg | 195 | 200 | 624 |
| | 200 | 205 | |
| gcg tat gga aca gac ccg aca cta aag ttc acg cag cac att cag agc Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser | 210 | 215 | 672 |
| | 215 | 220 | |
| aag gtg aag ccc gag gtg att gtg aaa atc tac aag gat ttc gtg att Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile | 225 | 230 | 720 |
| | 230 | 235 | 240 |
| ccg ctc acg aag gtc gaa gtc gac tac ttg ctg aga cgg ctg gag Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu | 245 | 250 | 768 |
| | 250 | 255 | |
| gac gag gag gac gat gat gcg acg cag aaa agc ggc ggc tac gtt gac Asp Glu Glu Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp | 260 | 265 | 816 |
| | 265 | 270 | |
| cggtttctctccatggtttacatgt Arg Phe Leu Ser Ser Gly Leu Tyr | 275 | 280 | 843 |
| | 280 | | |
| <210> 2 | | | |
| <211> 280 | | | |
| <212> PRT | | | |
| <213> Hansenula polymorpha | | | |
| <400> 2 | | | |
| Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg | | | |
| 1 | 5 | 10 | 15 |
| Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu | 20 | 25 | 30 |
| Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe | 35 | 40 | 45 |

Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln
50 55 60

His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu
65 70 75 80

Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile
85 90 95

Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn
100 105 110

Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala
115 120 125

Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala
130 135 140

Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg
145 150 155 160

Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp
165 170 175

Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn
180 185 190

Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg
195 200 205

Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser
210 215 220

Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile
225 230 235 240

Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu
245 250 255

Asp Glu Glu Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp
260 265 270

Arg Phe Leu Ser Ser Gly Leu Tyr
275 280

<210> 3
<211> 1655

<212> DNA
<213> Hansenula polymorpha

<220>
<221> gene
<222> (1)..(1655)
<223> 1,8 kb genomic DNA-fragment from Hansenula polymorpha

<220>
<221> gene
<222> (1)..(1655)
<223> 1,8 kb genomic DNA-fragment from Hansenula polymorpha

<400> 3
cccgccccaa tgccagcaat atggagacgt ttaggcagaa taggcgttcc atacttctca 60
cgctgcttgt tgccaccgga atatacacccg cattgcagtt tgcacacatc atactatatg 120
acgattacat tggcggAACG tatcgcgagt cgctcacgag acgcattaga atgacagaga 180
aatcgcgaaa cgaccttata gacgcacgtg aaaactacgg gtttggaggc agcaaggagg 240
agcgaatcca gcggttttg tggtcagac atcttcgtg gcttttaggc gaggataagc 300
gaacttgagg agcgTTTTT tttcctgtt tagTTTTGT aggtatggac tttatgaagc 360
cagaaacagt gctggacctt ggcaacatta gagatgcctt ggtccggatg gaggataacga 420
tcatcttcaa ctttatcgag cggtcgcaGT tctatgcgtc gccctcggtt tacaaggatca 480
accagttccc tattccaaac ttgcacggct cgTTTGGt ctggctgtt tcgcagcacg 540
agcgaatcca ttgcaggtg aggagatacg acgcGCCAGA cgaggtgcct ttttccccca 600
acgtgctgga aaaaacgttt ctgcccaga tcaactaccc atcggtgcta gcctcctacg 660
cgatgaaat caacgtcaac aaagagatac tcaagatcta cacgtcagag atagtaccag 720
gaatagctgc aggcaGCGGA gaggcaggagg acaaccttgg ctcgtgcga atggccgaca 780
tcgagtgcct gcagtcgcta tccagaagaa tccatTTGG ccgtttgtc gcagaggcta 840
aatttatcag tgagggggac aagattgtgg atctgatcaa aaagagagat gtggaggca 900
ttgaggcgct catcacaaac gccgaggtcg aaaaacggat cttggacaga cttctggaga 960
agggaggGC gtatgaaaca gacccgacac taaagttcac gcagcacatt cagagcaagg 1020
tgaagcccga ggtgattgtg aaaaatctaca aggattcgt gattccgctc acgaagaagg 1080
tcaagtcga ctacttgctg agacggctgg aggacgagga ggacgatgat ggcacgcaga 1140
aaagcggcgg ctacgttgac cggttctct cctctggctt gtactagaaa taaaatttt 1200
cagtacttta attattctcg aattcttagtt cagataccgc atggtaattt caaaggccag 1260
aaaagtggcc gcgttggctg gggcagctct cagaatagtc ggcgagaatc ctttgactag 1320
cccccaggca ccgctctgtc tccaaatacc cctaatagtc tcaacagcat ttctataaac 1380
cagcttcttg tagttgtccg tctgcattttt ggacttgcac acatcgatcg gataaataact 1440

gaaccacatc ccgtaacctg ccagcgcccc aaagacgcag agcttccagt tctcgatgtc 1500
cttcctggca atattccgcg actcgatctc gttttcacg agagcttcaa aagtcagaaa 1560
atacgctccg ctacccaaac tttctttgc cagcgttaggt cccagacccc ggtagattaa 1620
cttgatgcct cccgtatggt acagcttctt gatcc 1655

<210> 4
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 4
aattaaccct cactaaaggg

20

<210> 5
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 5
gtaatacgac tcactatagg gc

22

<210> 6
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 6
atatacatct acaaaaacta aacagg

26

<210> 7
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 7
atatacatct gatgcgacgc agaaaagc

28